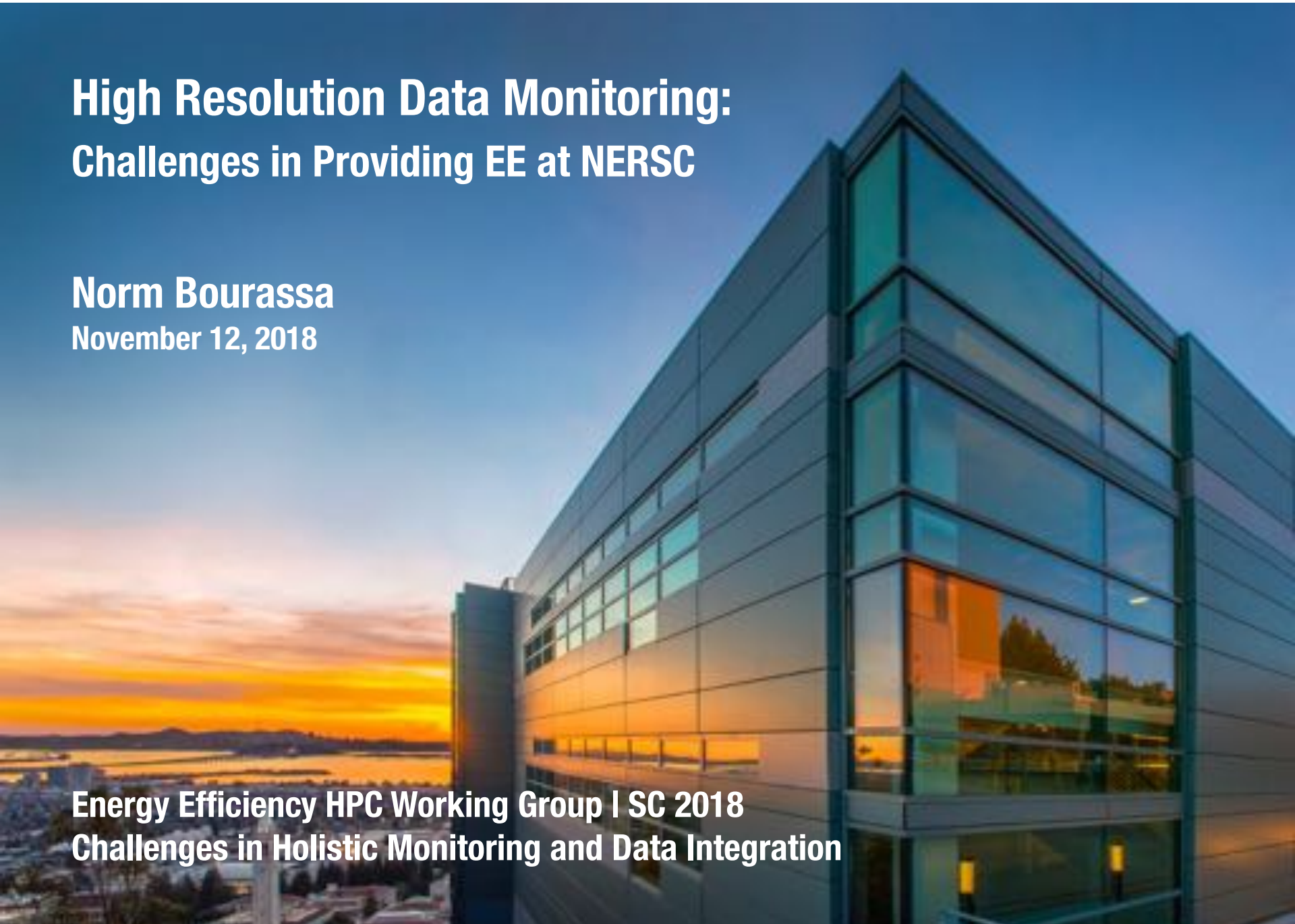


High Resolution Data Monitoring:
Challenges in Providing EE at NERSC

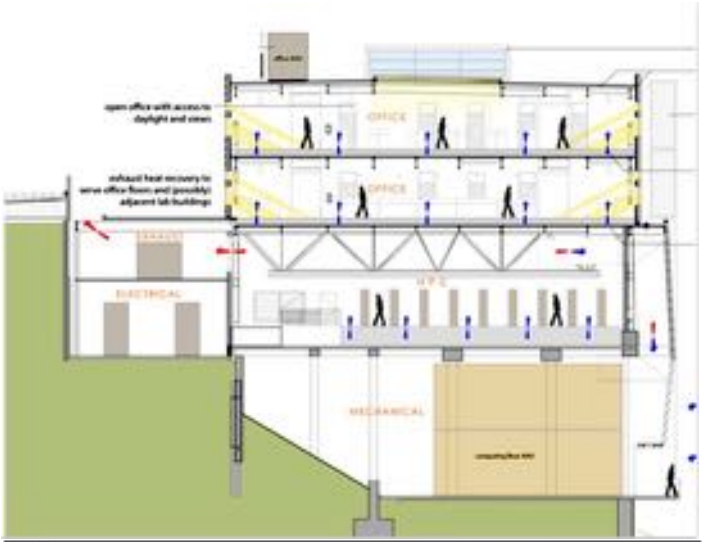
Norm Bourassa

November 12, 2018

Energy Efficiency HPC Working Group | SC 2018
Challenges in Holistic Monitoring and Data Integration

Shyh Wang Hall (B59) completed June 2015

- **Four story 150,000 GSF**
  - Two 20ksf office floors, 300 offices
  - 20k -> 28ksf HPC floor
  - 12.5MW expandable to 20+MW
  - Current demand is ~7MW
- **Energy Efficient**
  - Year-round compressor free air and water cooling
  - LEED Gold
  - **PUE < 1.1**
- **Room Air & Cooling Water**
  - Compression free cooling
  - Most hours, single pass air
    - Humidity Control very challenging
  - Cooling Towers can provide 75°F or cooler water all hours of year



General Update

- **Cori entered production on July 1, 2017**
  - Compute intensive and data intensive workloads
  - 9688 KNL nodes; 2388 Haswell nodes
  - NVRAM Burst Buffer 1.5PB @ 1.5 TB/s
  - 30 PB Lustre, > 700GB/sec
  - SDN for enhanced external networking connectivity
  - Containerized software with Shifter

- **Recent Projects with Data Collect**
  - Function as a Service deployment (FaaS)
  - Real-time PUE calculation
  - B59 significant focus of LBNL campus EE Program
    - Multiple projects completed and in progress
  - Cray XC Series fan control monitoring
  - Hot Aisle Containment improvements

Data Collection System

- **Current Data Sources**
  - Substations, panels, PDUs, UPS
  - Cori & Edison SEDC
  - One-wire Temp & RH sensors
  - BMS through BACNET
  - Indoor & Outdoor Particle counters
  - Weather station

- **Future Data Sources**
  - Syslog
  - Job Data
  - Lustre + GFPS statistics
  - LDMS
  - ???

- **Rabbit MQ, Elastic, Linux**
  - Collects ~20K data items per second
  - Over 40TB data online (100TB capacity)
  - 45 days of SEDC (versus 3 hours on SMW)
  - 180 days of BMS data (6X more than BMS)
  - 260 days of power data

Kibana, Grafana

Data – User Access & Public Security Tension

- **Challenge: General Cyber security vs User direct access to monitoring data streams**
  - UI tools (Grafana, Kibana) provide visual access only
- **Using Fn Project project to provide User "on demand" access to gathered data**
  - FaaS – Function as a Service
  - Provides REST-like API for users to link their applications to NERSC Elastic data

PUE Measurement and Reporting

- **Realtime PUE**
  - Monthly average has been ~1.08 (May – Sept., 2018)
  - Calculating Level 1 and 2 PUE
    - Difference never > 0.01
  - Challenge: ~ 60 measurement points at differing measurement intervals
  - Testing several calculation methods
    - 15-minute interval calculation – Many fails since substation metering is at same interval rate
    - Last value 5-minute rolling window – Non-coincident data points, but produces smoother results
    - Moving average 5-minute window within 30-minute sample captures – Measurement averaging helps reduce effect of non-coincident data points in 5-minute rolling window alone

PUE Measurement and Reporting

- **50+ Monitoring Points**
  - High granularity for LBNL Campus "Mechanical Metrics" within SkySpark energy info system

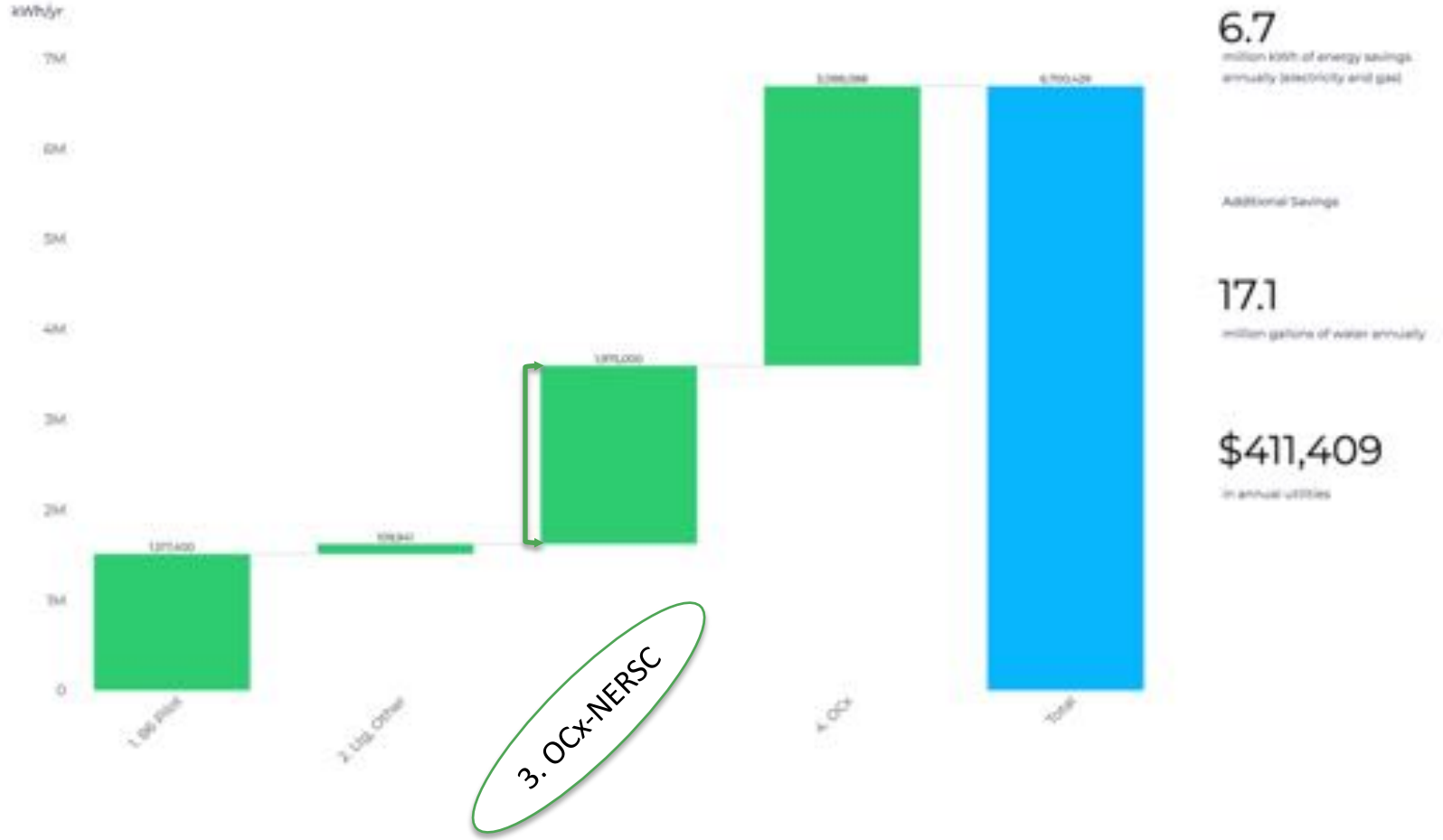
PUE Calc. Point	Substation	PUE Level	Description
A	Building	1 & 2	INCOMING MAIN #2
B	590	1 & 2	MVSG SUBSTATION 590
B'	590	1 & 2	DISTRIBUTION TRANSFORMER 590 OUTPUT
C	596	1 & 2	MVSG SUBSTATION 596
C'	596	1 & 2	DISTRIBUTION TRANSFORMER 596 OUTPUT
D	612	1 & 2	MVSG SUBSTATION 612
D'	612	1 & 2	DISTRIBUTION TRANSFORMER 612 OUTPUT
E	613	1 & 2	MVSG SUBSTATION 613
E'	613	1 & 2	DISTRIBUTION TRANSFORMER 613 OUTPUT
F	628	1 & 2	MVSG SUBSTATION 628
F'	628	1 & 2	DISTRIBUTION TRANSFORMER 628 OUTPUT
N1	590	1 & 2	COMPUTE LIGHTING
N2	590	1 & 2	GENERATOR HEATER, BATTERY CHARGER,
N3	590	1 & 2	VESDAS, PREACTION AIR COMPRESSOR
N4	590	1 & 2	590A5A PANEL: AHU LTG/PWR. RECPTACLES
N5	590	1 & 2	590A6A PANEL: DOLPHIN SYSTEM, LTG.
N6	590	1 & 2	E590A7A PANEL: AHU-001/11 (COMPUTE), AC-
N7'	590	1 & 2	UE600A ENTIRE PANEL, TO BE USED TO
N7"	590	1 & 2	OFFICE (NON-COMPUTE) LOADS ON UE600A
N8	590	1 & 2	595A40A PANEL: COMPUTE ROOM 2101
N9	590	1 & 2	COMPUTE LIGHTING, 590A15A
N10'	596	1 & 2	596A1A1A ENTIRE PANEL, TO BE USED TO
N10	596	1 & 2	ELEVATOR EL-01 (FREIGHT), 596A1A1A39A
N11"	596	1 & 2	OFFICE LOADS, 596A1A2A
N12"	596	1 & 2	Office Space Heat Loads, 596A1A3A

PUE Calc. Point	Substation	PUE Level	Description
ND1-1	590	1	UPS-1
ND1-2	590	1	UPS-2
ND1-3	590	1	PDU 10
ND1-4	590	1	PDU 12
ND1-5	590	1	PDUs 26, 28,
ND1-6	590	1	PDUs 22, 24,
ND2-1	590	2	UPS PDU
ND2-2	590	2	UPS PDU
ND2-3	590	2	UPS PDU
ND2-4	590	2	UPS PDU
ND2-5	590	2	UPS PDU
ND2-6	590	2	UPS PDU
ND2-7	590	2	UPS PDU
ND2-8	590	2	UPS PDU
ND2-9	590	2	EDISION
ND2-10	590	2	CORI PDU
ND2-11	590	2	HOUSE PDU
ND2-12	590	2	HOUSE PDU
ND2-13	590	2	HOUSE PDU
ND2-14	590	2	HOUSE PDU
ND2-15	590	2	HOUSE PDU
ND2-16	590	2	HOUSE PDU
ND2-17	590	2	HOUSE PDU
ND2-18	590	2	HOUSE PDU

PUE Measurement and Reporting

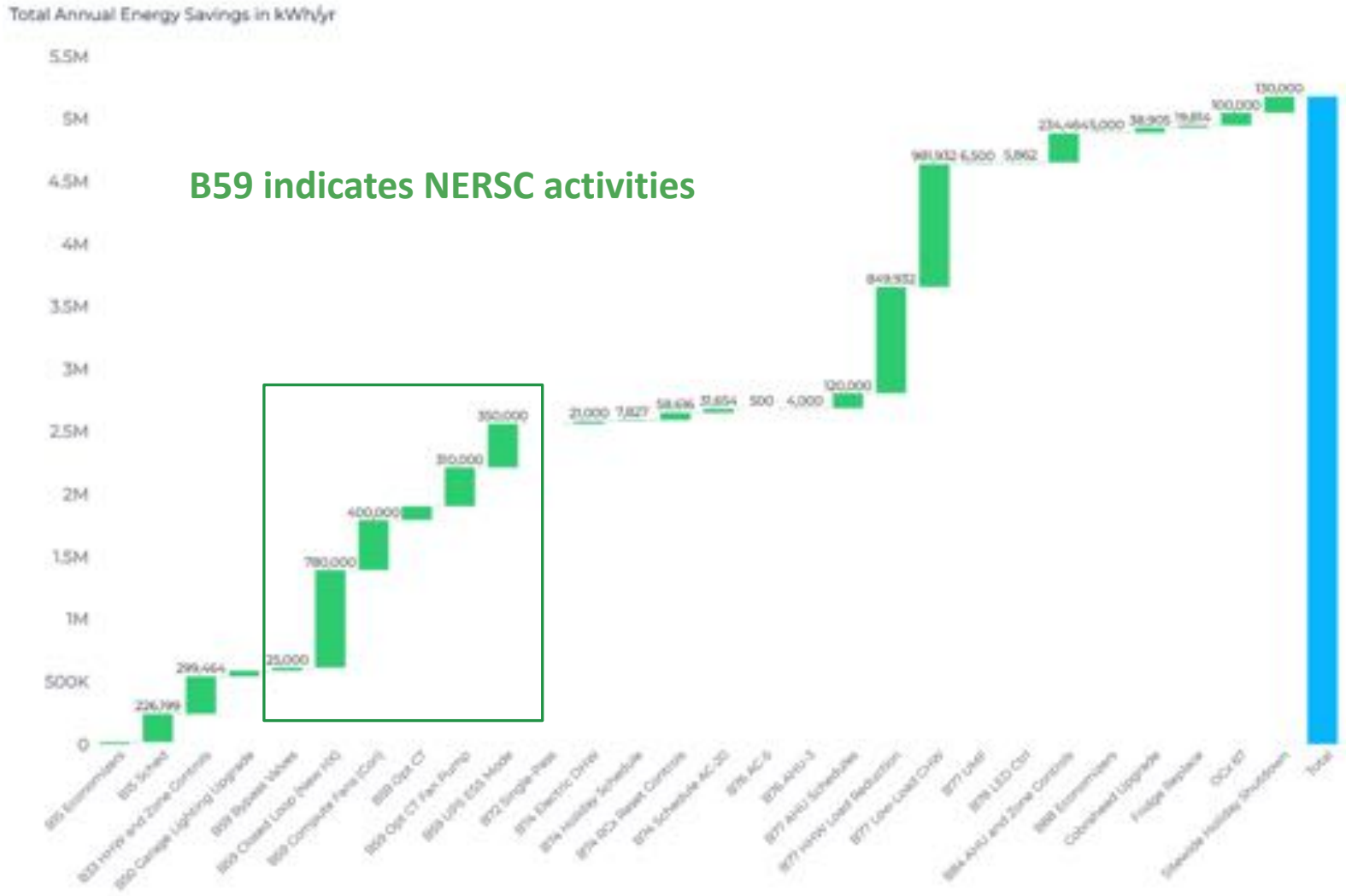
LBNL Campus Wide Efficiency Savings

The LBNL Campus is able to generate significant, maintained savings - primarily from improvements in building operations



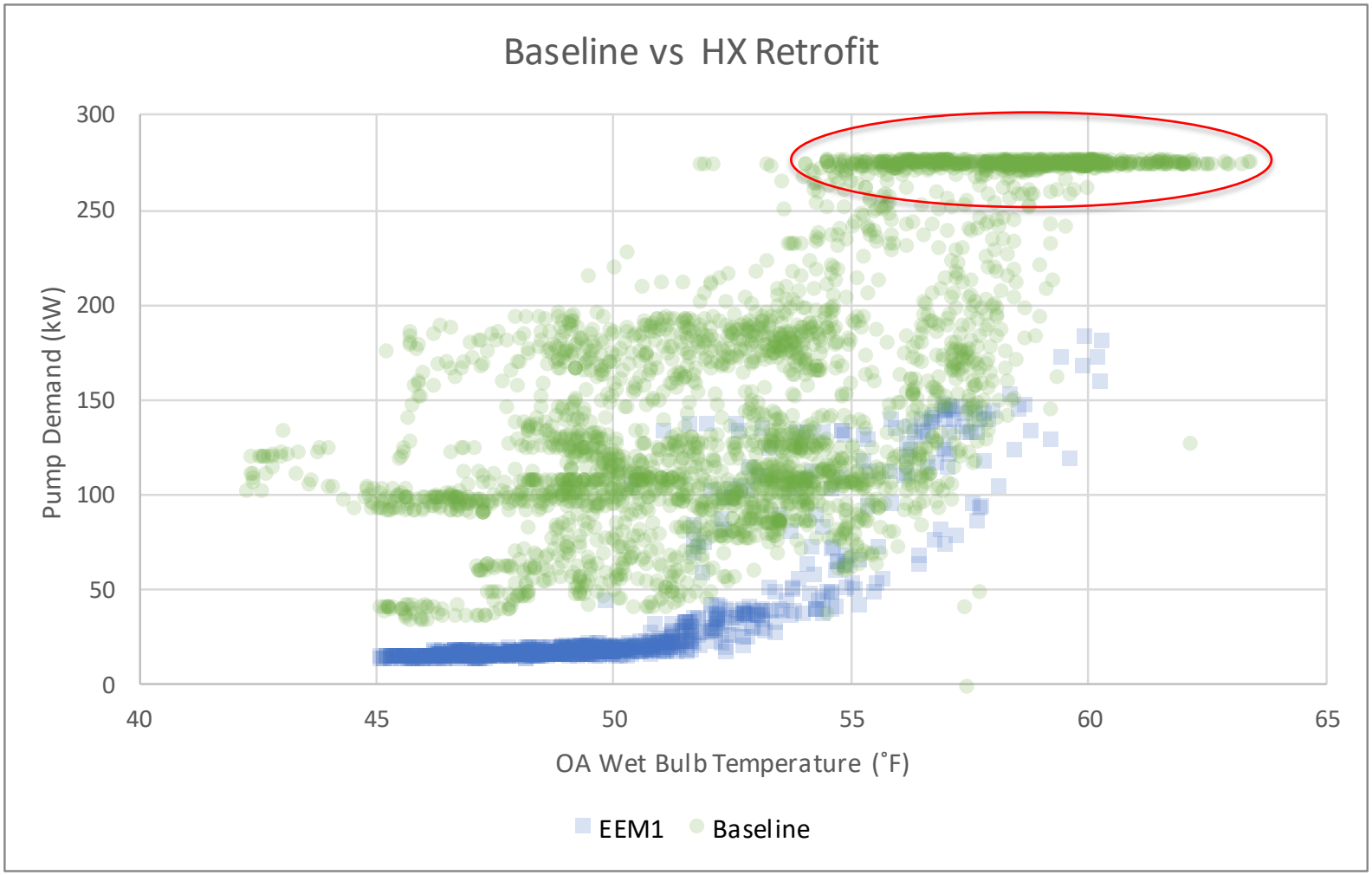
Ongoing Commissioning

LBNL Wide - Ongoing Commissioning

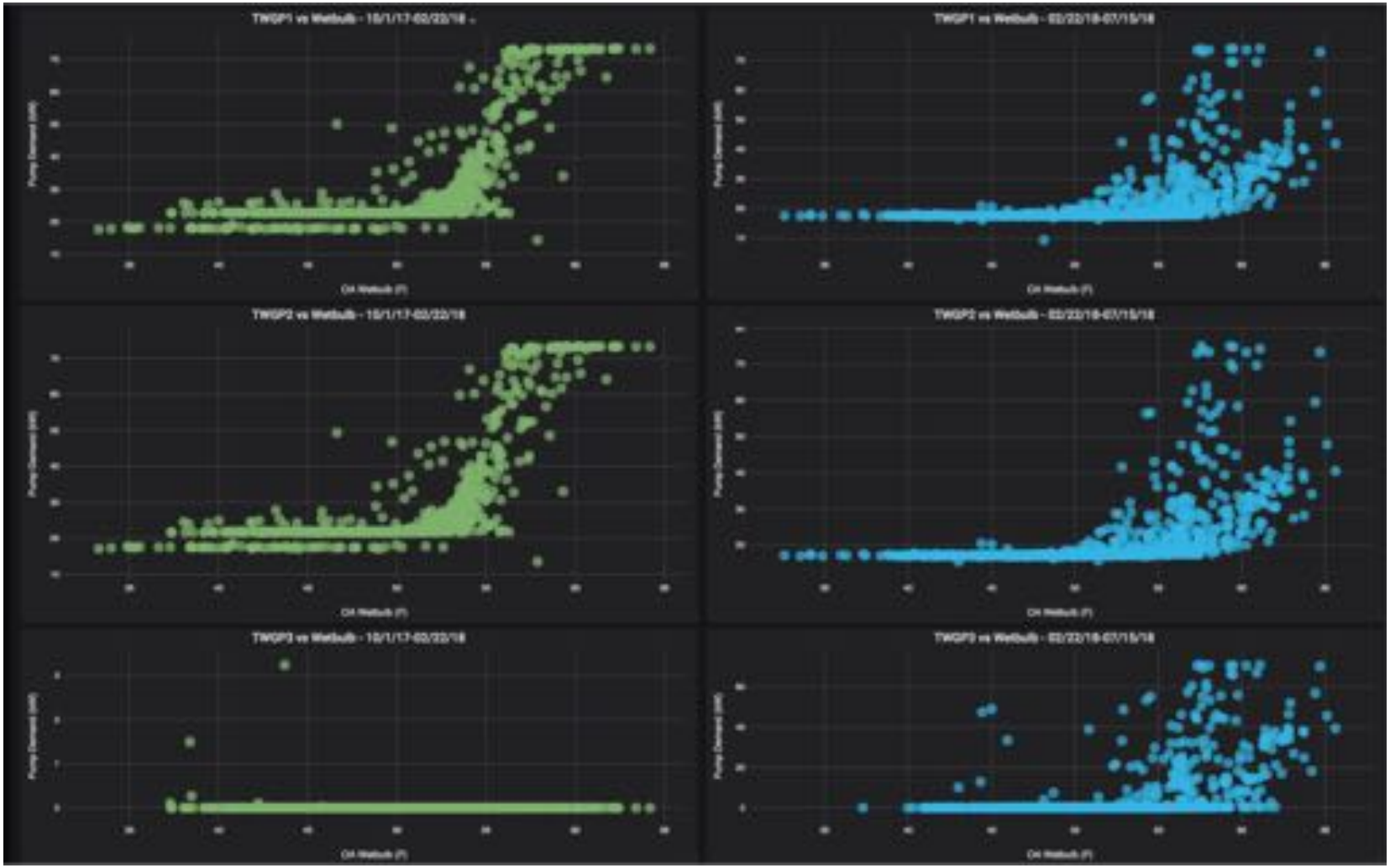


Notes:

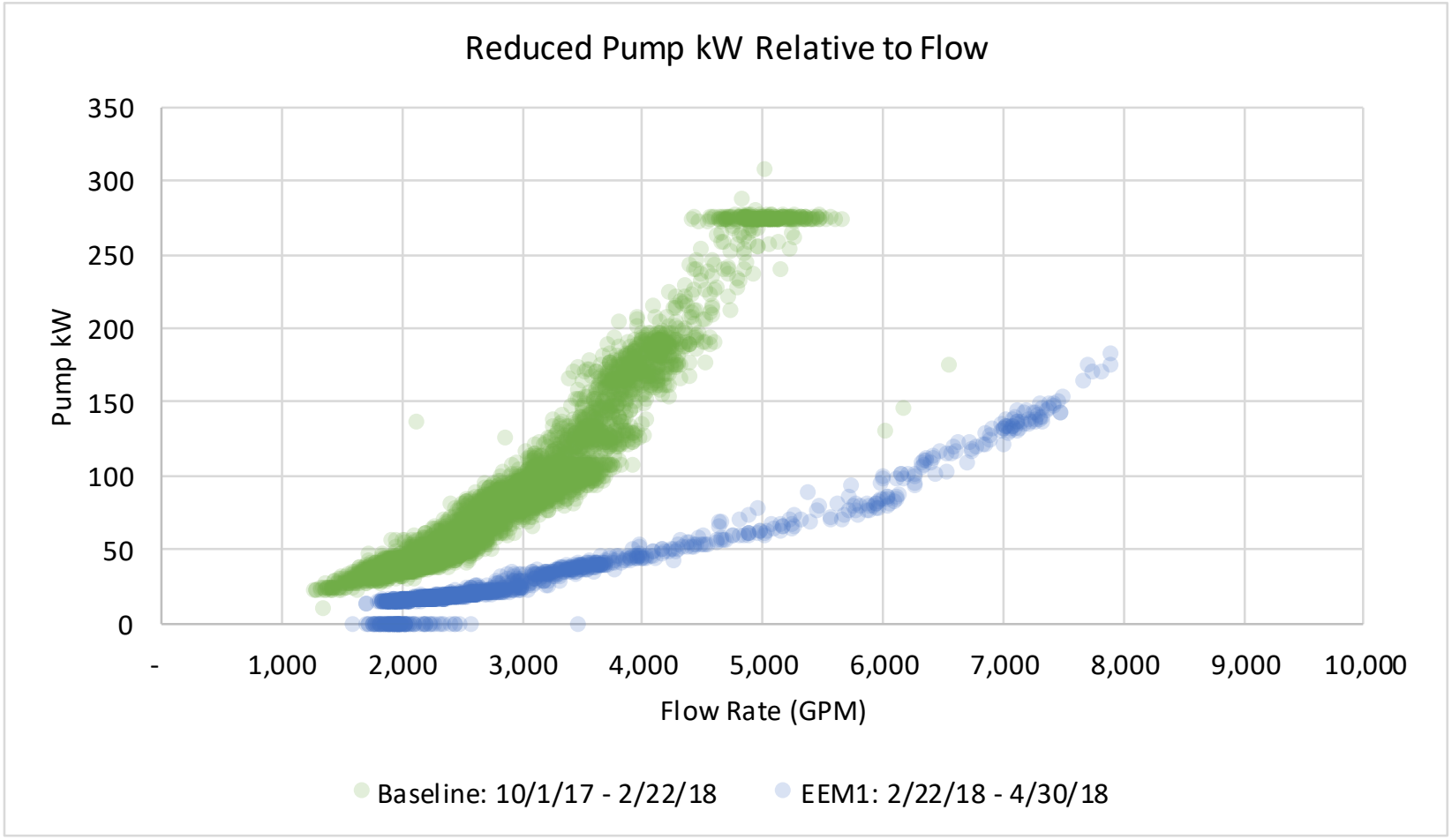
Additional CW Heat Exchanger & Pump Installed



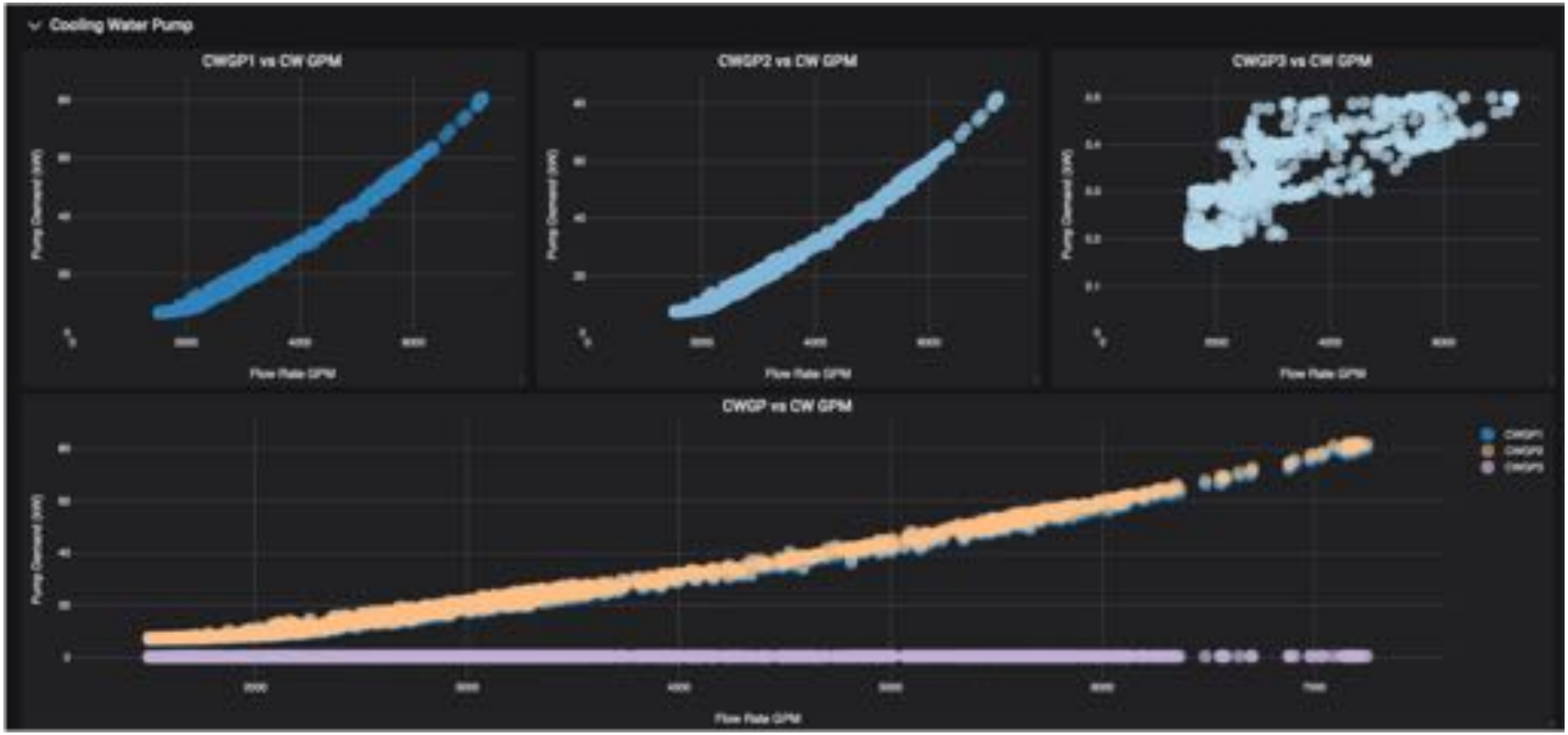
In Development – Live data Scatter Plots



Additional CW Heat Exchanger & Pump Installed

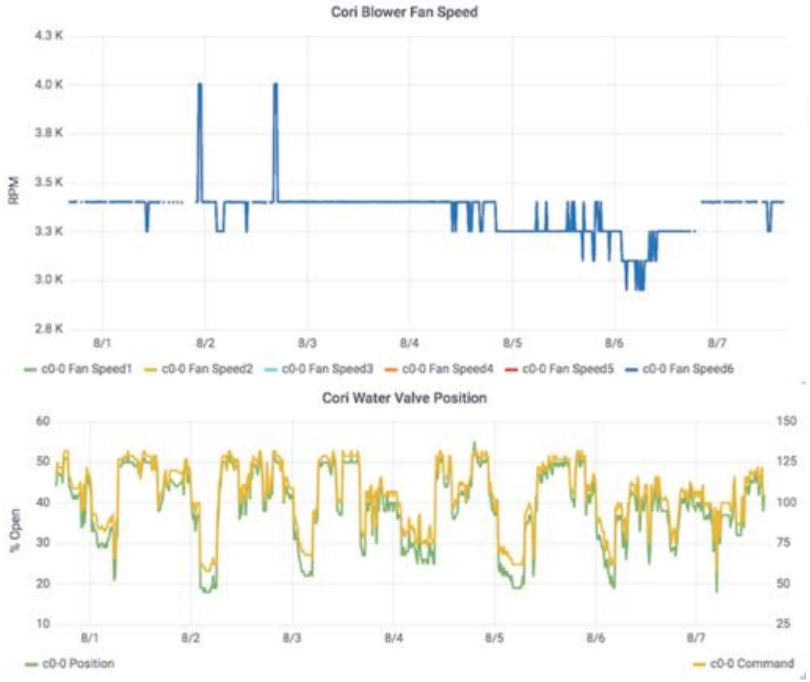


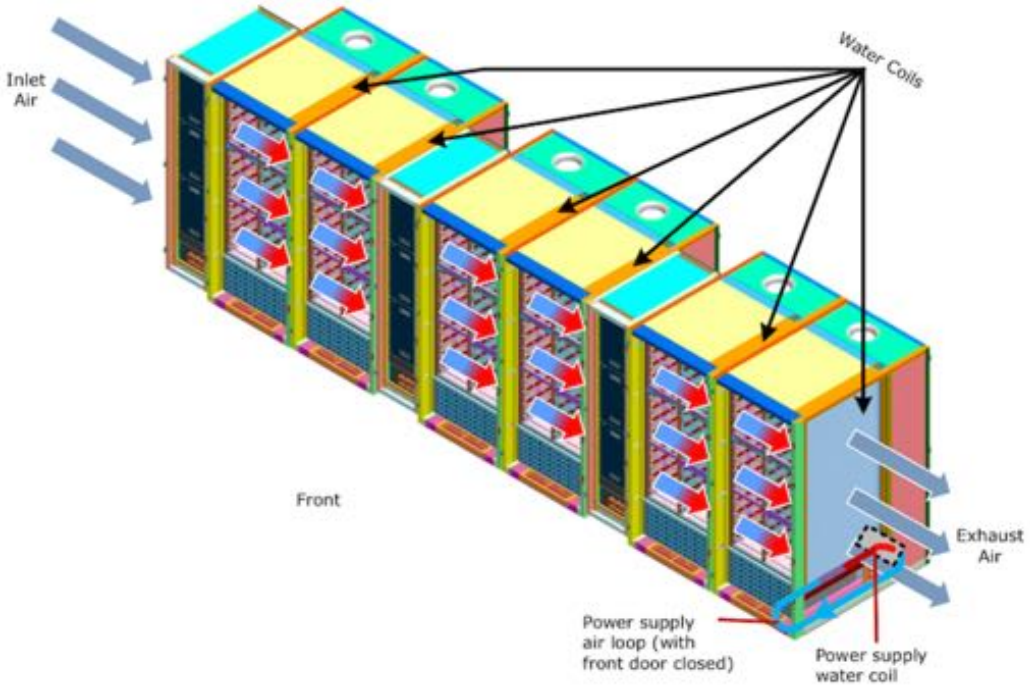
Recent Development – Live data Scatter Plots



Cray XC Series - Dynamic Fan Speed Control

- **Blower Cabinets (6 High Performance Fans per)**
  - Variable from 2500 to 4000 rpm
    - Dynamic Fan Speed Control saves energy by sensing 5°C processor temp change (HPC load), then
    - Adjusts fan speed (entire row) up or down accordingly. Hottest processor(s) set speed on entire row
  - Cooling Water Supply Temp control of cooling coil exiting air temp. (Typ. 68 – 71°F)
    - Better transfer of cooling load onto CW plant when water cooling conditions favorable





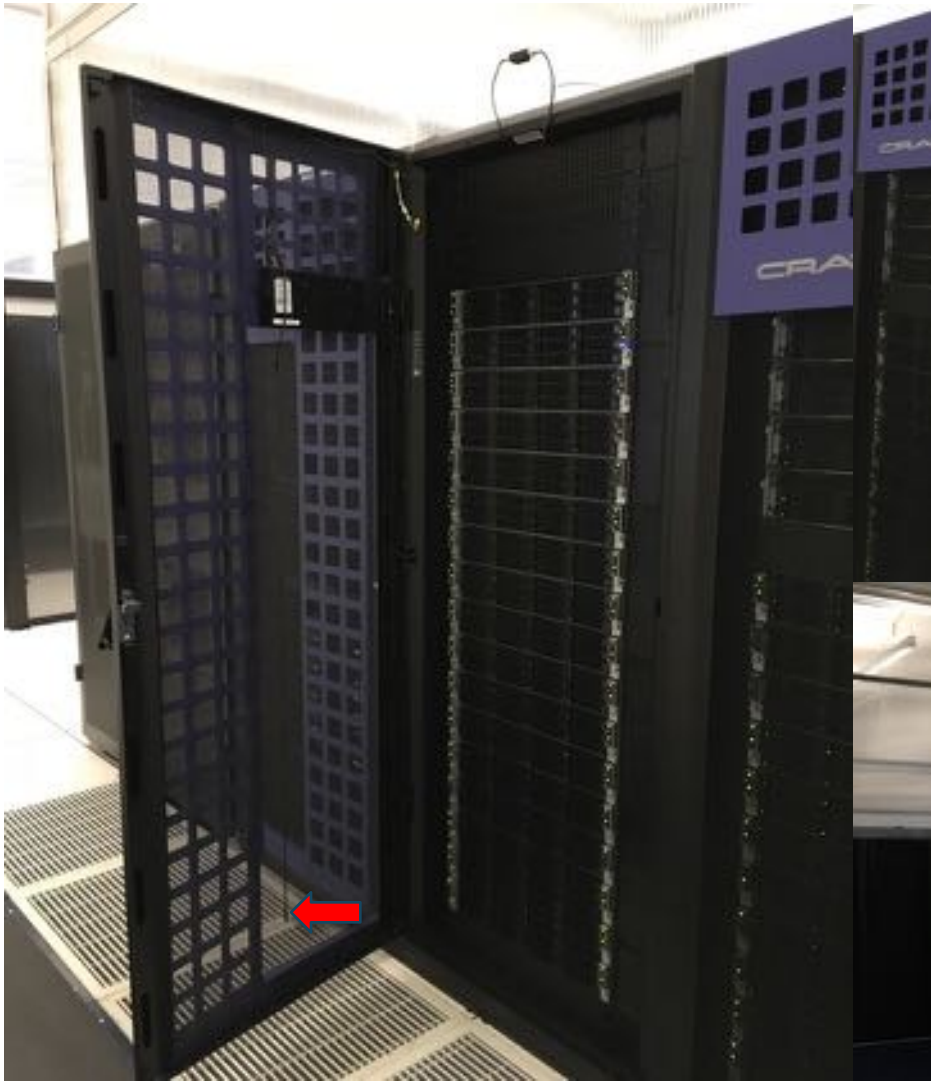
Cray XC Series - Dynamic Fan Speed Control

- **April 6, 2018 - DFSC turned on for all of Cori**
  - Non-Dynamic defaults are 2 fan speeds
    - Normal = 3400 RPM; Max = 4000 RPM
  - Dynamic Fan Speed Control multiple fan speeds (user defined, max 15 steps)
    - 150 RPM per 5°C processor temp change; Minimum = 2500 RPM; Max = 4000 RPM
    - 11 Steps at NERSC – Typical speeds are #6 and 7. Rarely above #8.

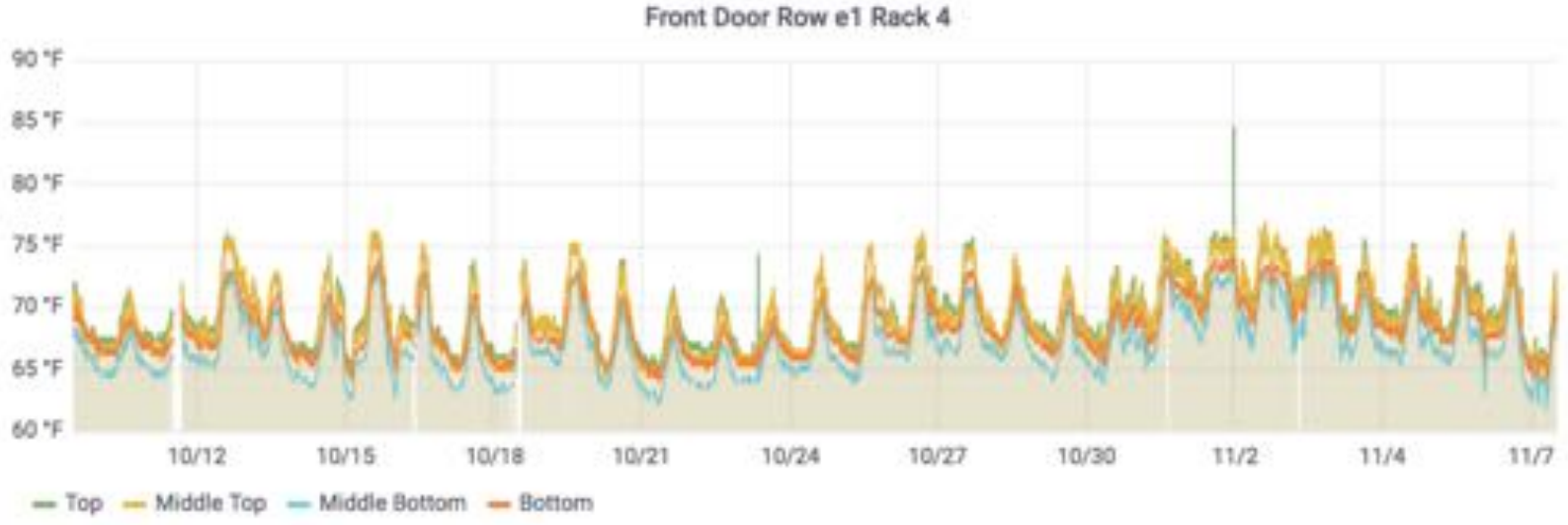
Table Step	Fan Speed (RPM)
11	4000
10	3850
9	3700
8	3550
7	3400
6	3250
5	3100
4	2950
3	2800
2	2650
1	2500

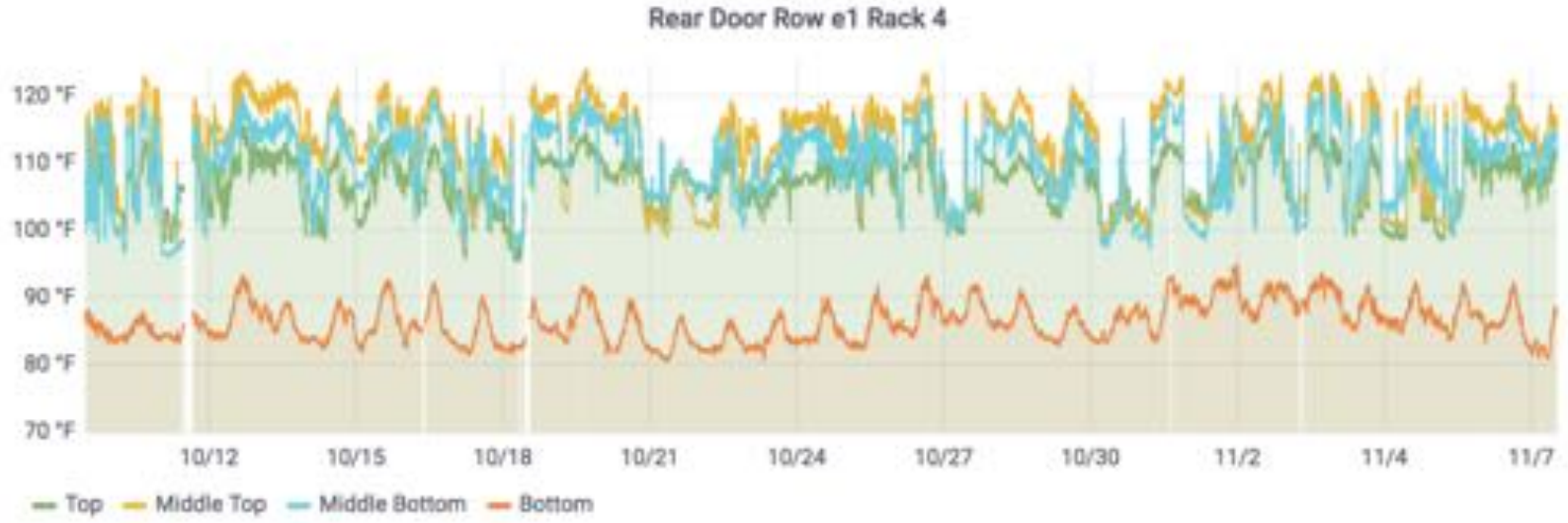
Hot Aisle Containment – Or Hot Air ”Bath Tub”

- **Really is a “separation” or “chimney” system, not full containment**
  - Installed in stages from 2016 -2018
  - Panelized system better accommodates equipment churn
- **Helps drive hot air to ceiling**
  - Exhaust fans pull hot air out of the building
  - Return fans to AHUs
- **Panels located at rack front**
  - Provides cabling facilitation above rack without penetrations
- **Stayed away from drop ceiling**
  - Seismic floor connection costs
  - Fire code complications
- **Why Hot instead Cold Aisle?**
  - Many visitor tours of compute room, so human comfort a factor

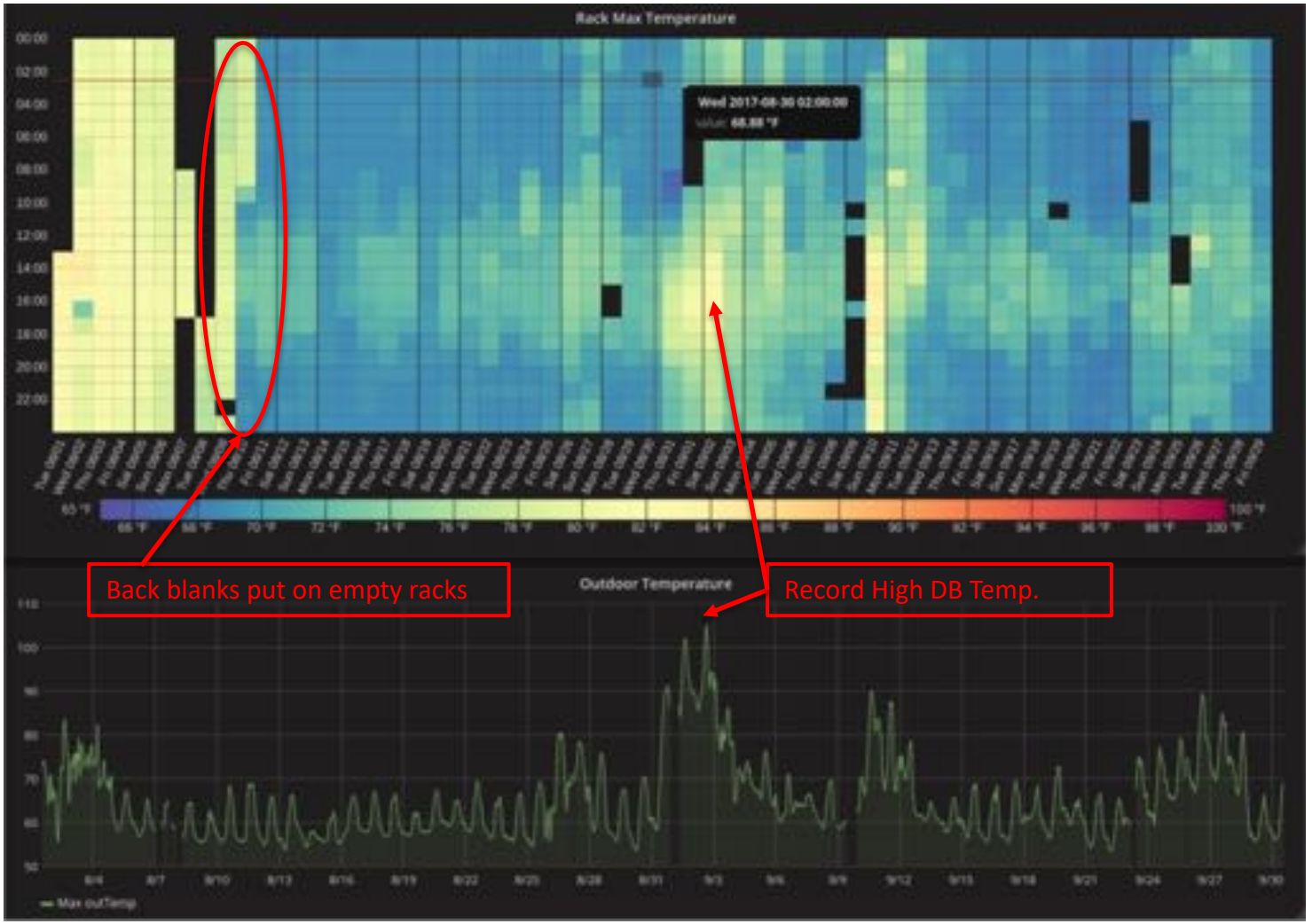


Potential night time over supply

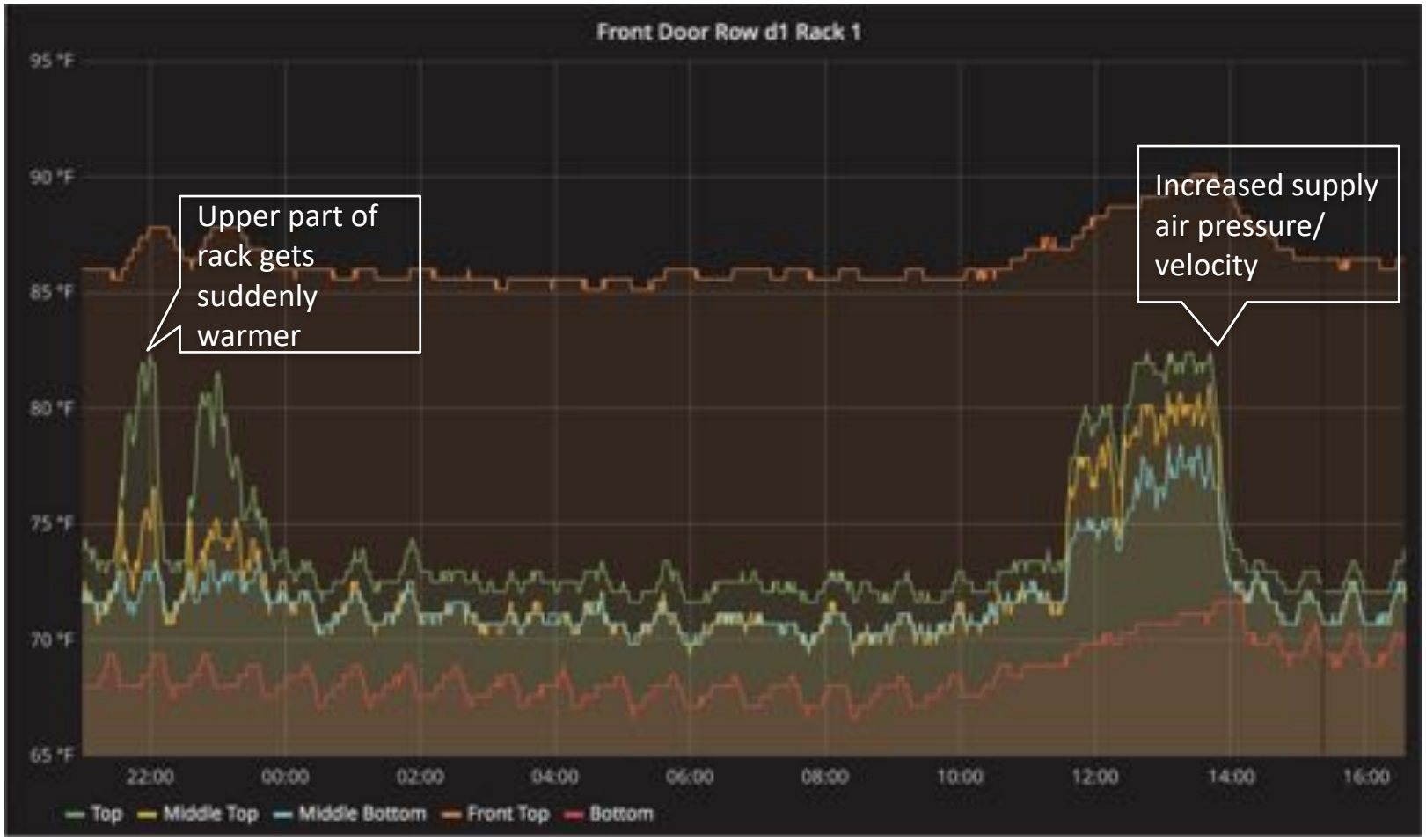




Data Helped identify air sealing measures



Bath Tub Effect Breaks



Data Collect Challenges

- Data analysis can be difficult – Good visualization tools a necessity
  - Good problem to have: Access to more raw data than is needed for task at hand
- Native Elastic Data UI Tools are limited (but improving)
  - Requires a level of data coding aptitude from users
    - User can easily create data queries (accidental or not, thus hanging entire system
    - A challenge for many staff & upper management
  - Visualization pallet is limited
  - Obtaining native sample resolution is difficult
    - UI Tools auto interpolate data interval based on time range being viewed
- Current system requires considerable admin attention
- Public security vs user data access a perpetual challenge
- Obtaining BMS control grade from native Elastic sensors is expensive

National Energy Research Scientific Computing Center

Norm Bourassa, njbourassa@lbl.gov

Measures Overview

		Energy Savings (kWh)		Water Savings	Cost Savings	PUE
	Measure Title	Estimated	Verified	Gallons	$	Reduction
1	Install Firmware to Enable ESS Mode for UPSs		350,000	140,000	$20,300	0.007
2	Implement Tower Water Supply Temperature Reset and Reduced Tower Water Pump Speed		420,000	170,000	$24,360	0.009
3	Reset Cooling Water Temperature Setpoint and Enable Cray Dynamic Fan Control		400,000	160,000	$23,200	0.008
4	Install New Heat Exchanger		780,000	310,000	$45,240	0.016
5	Install Bypass Valves		25,000	10,000	$1,450	0.001
6	Reset Cray Air Temperature Setpoint	200,000		80,000	$11,600	-
7	Optimize Dynamic Fan Control	200,000		80,000	$11,600	-
8	Install Booster Pump	240,000		100,000	$13,920	0.005
9	Install Cold Aisle Temperature Sensors and Optimize AHU SAT and Flow Control	300,000	-	120,000	$17,400	0.006
10	Install Cray Supply Air Hoods	100,000	-	40,000	$5,800	0.002
	Total	**1,040,000**	**1,975,000**	**1,210,000**	**174,870**	**0.054**

Sept. 2018: PUE = 1.08 (avg.)

U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB